Lexicalized ontology for the management of business rules
An industrial experiment

Nouha Omrane* Adeline Nazarenko* Peter Rosina Sylvie Szulman* Christoph Westphal****

*LIPN UMR 7030 (Université Paris 13 & CNRS), France

`firstname.lastname@lipn.univ-paris13.fr`

** AUDI AG, Germany

`firstname.lastname@audi.de`

Abstract

This paper[1] describes a simple formalism designed to encode lexicalized ontologies and shows how it is used in a business rule management platform[2] of the automotive domain.

1 Introduction

Business Rules Management Systems (BRMSs) are software applications that help organizations to separate their application code from their business knowledge. BRMSs help the users to author and maintain business rules and apply decision logic that reflects this business knowledge. However, domain experts who are not always business rules experts may have difficulties expressing their knowledge in formalized logic languages. Supporting them in their management of the knowledge needed to write these rules is one of the goals of the ONTORULE project.

We propose building an ontology as a formal model for representing conceptual vocabulary that is used to express business rules in written policies. OWL-DL language is used to represent concepts and properties of the domain ontology but such an ontology must be linked to the lexicon used to express rules in the text, so experts can query source documents. This calls for a formalism to link linguistic elements to conceptual ones. We opt to use the SKOS[3] language which provides basic elements to link domain concepts to terms from the text. The combination of OWL entities, SKOS concepts and their related information form a lexicalized ontology which supports the semantic annotation of documents.

The paper is organized as follows. Section 2 describes the Audi use case, on which this approach has been tested. Section 3 explains the choice of SKOS combined with OWL as language to support the lexicalized ontology. Section 4 reports the experimentations made in the Audi use case.

2 The Audi use case

Nowadays, the development of new cars has become very challenging and many different process steps are involved. Computer Aided technologies, like virtual modeling, simulations or the analysis and planning of physical testing, need to be integrated even tighter to satisfy the higher requirements and reduced time-to-market which also shortens the development cycles.

In the ONTORULE project, Audi is developing a prototype BRMS that makes use of ontologies and business rules. Ontologies together with business rules help Audi to keep abreast of technology advances and use them in its R&D IT applications. Especially the interweaving of the various Computer Aided technologies will help Audi to reduce development time and cost.

One of the difficulties with business knowledge rules is that various departments or roles sometimes use different vocabularies for the same things so they cannot understand each other immediately. Addionally, formalized rules *per se* are often not easy to understand. Using an ontology as a unified model for a heterogeneous vocabulary will reduce misunderstandings and ensure that people are discussing the same thing. Also, the users can easily confirm and verify the appropriateness of

[1] This paper is an extract from our paper (Omrane et al., 2011a)

[2] This work was realised as part of the FP7 231875 ONTORULE project (http://ontorule-project.eu). We thank our partners for the fruitful discussions, especially to Audi for the collaboration on their use case.

[3] Simple Knowledge Organization System

the modeled semantic relations. Finally, the prototype that is to be developed is expected to handle links between source documents, such as policies or internal documents, and the concepts and instances of the ontology.

3 A formalism for lexicalized ontologies

3.1 Existing formalisms

Many research activities have tackled the problem of linking an ontology to a lexicon. Two major areas are of interest. The first is the NLP domain which aims at adding some semantic structure to a lexicon by linking its elements to ontology's elements. There are several ways to combine a lexicon with an ontology: *LMF*[4] standard (Francopoulo et al., 2007), *TMF*[5], *OLIF*[6], *LMM*[7]. The other group trys to link an ontology to a lexicon by modeling linguistic information in the ontology as in (Reymonet et al., 2007), *LexOnto* (Cimiano et al., 2007) or *LIR* (Peters et al., 2009). There also exist more abstract approaches like *LingInfo* (Buitelaar et al., 2006), which defines a meta class to link the linguistic properties to the concept or to its Data/Object properties, or (Ma et al., 2009), which introduces a set of annotation rules to link an existing ontology to its lexicon.

From a practical point of view, the choice of one model or another depends on the aimed application and the task. Our aim is to build a lexicalized ontology to allow annotating the technical documents and thus to help the expert in exploring documents by querying its set of annotations. We use for that *SKOS* W3C standard that links linguistic to semantic knowledge.

3.2 A SKOS-based approach

A key issue for experts in managing a rule base is to be able to mine textual sources to understand how a given concept is used in business documents, what rules are related to it and how those concepts and rules evolve when the policies are updated. This is achieved through the semantic annotation of the documents in which the mentions of the ontological entities (concepts, instances and roles) are highlighted and can be searched for.

[4]Lexical Markup Framework

[5]Terminological Markup Framework

[6]Open Lexicon Interchange Format

[7]Linguistic Meta Model

Our aim is therefore to save the terms related to the conceptual vocabulary that is used to express the business rules. We don't need to encode sophisticated information such as the morphological structure of terms since we do not perform a deep analysis of the documents. We simply need to save the various linguistic units that denote a concept, instance or role. SKOS supports encoding of *SKOS concepts* that represent the links between the OWL concepts and their related terms, which are encoded as *skos labels*[8]. This relation is described by `<rdf:Description rdf:about>`.

When designing and updating business rules, experts face the problem of the heterogeneity of information sources and multilingualism. *SKOS* also supports that normalization of vocabularies. A given SKOS concept can be associated with the various terms or labels that denote it in the texts or any other information source. For a given concept, *SKOS* supports distinguishing one preferred label and as many alternative labels as necessary, using the `<skos:prefLabel>` and `<skos:altLabel>` properties. In the Audi ontology, for example, the SKOS concept *LowTemperatureChamber* is linked to two terms: *low temperature chamber* is encoded as the preferred label and *refrigerated cabinet* as its alternative form.

SKOS also supports the encoding of multilingual information. The information about the language used is described by `<rdf:lang=''en''>`. For example, the SKOS concept *TrolleyTest* has a preferred label "trolley test" which is mentioned in English texts, and an alternative label "Schlittentest" in German.

Since experts often have to manage a large volume of information but do not always formally describe all the concepts, it is important to add informal documentation when it is available. Defining concepts in natural language is very important to understand what concepts mean, especially if they have ambiguous or implicit labels. Those definitions can sometimes be extracted from the source documents when designing the ontology. In that case, they are associated to the related SKOS concepts using the label `<skos:definition>`.

In such a lexicalized ontology, the domain concepts and their occurrences in the text can be

[8]http://www.w3.org/2004/02/skos/

matched from one to another thanks to the linkage of OWL entities, SKOS concepts and labels. This is a simple efficient way to represent lexicalized ontologies and we show in the following section its benefit for the Audi BRMS. Figure 1 describes how the Audi ontology is linked to the lexicon and annotated text.

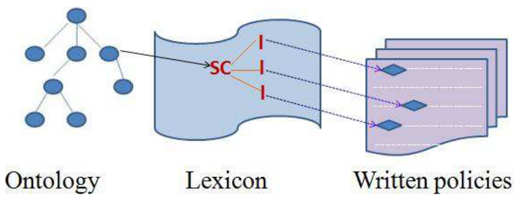

Figure 1: A lexicalized ontology for annotating source documents. Each concept from the ontology is linked to a SKOS concept *SC* and each SKOS concept is related to its labels. The annotations link some text entities to these labels

4 The Audi lexicalized ontology

This section presents the Audi ontology and illustrates the benefit in the Audi use case of having such a lexicalized ontology.

That ontology has been built in two steps. At first, the goal was to integrate the various existing knowledge sources in a single one. This resulted into a small conceptual model (around 30 concepts) associated with a large knowledge base (thousands of instances).

In a second step, in order to better fits the experts' needs for semantic querying and document mining, the initial ontology has been restructured and lexicalized. It also appeared useful to increase the granularity of the domain model so as to represent for instance not only the various types of tests but also their actual occurrences in the car manufacturing process (instances that are related to the different tests applied to specific vehicle models).

This led to encoding various elements as concepts rather than instances (90 concepts were added). The conceptual structure has been reorganized (4 subsumption levels instead of 1). A SKOS resource has been associated with this resulting ontology: each concept is related to at least 1 preferred label and up to 5 alternative labels. In addition, using a subset of the initial ontology for the exploration of written policies showed that some of the mentioned concepts were missing in the initial ontology and led us to enrich it (Omrane et al., 2011b).

Once the ontology is lexicalized, domain experts can query source documents to search for fragments of texts that describe specific concepts mentioned in rules. For example, they can find all references of the concept *BreakingStrengthOfStrapTest* in the text, wherever it is mentioned in the documents. They can also search for all sentences where the physical methods are mentioned in the text. As the concepts expressing tests are subconcepts of the concept "MethodInformation", we query the text by searching about all labels describing subconcepts of "MethodInformation".

Thanks to the labels of concepts, the ontology can be used to annotate the documents. Figure 2 shows an example of texts where all the mentions of known concepts are emphasized. This supports experts in browsing of documents.

Two belts or restraint systems are required for the buckle inspection, the low-temperature buckle test, the low-temperature test described in paragraph 7.5.4. below where necessary, the buckle durability test, the belt corrosion test, the retractor operating tests, the dynamic test and the buckle-opening test after the dynamic test. One of these two samples shall be used for the inspection of the belt or restraint system.

Figure 2: A fragment of text annotated by the lexicalized ontology.

5 Conclusion

The proposed integration of Computer Aided technologies will increase the flexibility of the development process, allowing Audi to meet the increasing market demand for product diversification. This integration relies on the design of an application that is currently under development and is based on a BRMS.

Our approach for the acquisition and management of the knowledge embodied in such BRMS relies on a lexicalized ontology which unifies and normalizes the various vocabularies and links the conceptual knowledge to the source policies and regulation written in natural language. Using a lexicalized ontology enables experts to determine the most suitable Computer Aided technologies from given functional requirements and to query sources documents.

These new approaches, standards and technologies are already partially integrated in some processes. During the next years Audi will continue to incorporate the ONTORULE platform in their landscape which will lead to even less timeconsuming, cheaper and higher quality processes in the innovation and development cycles.

References

Paul Buitelaar, Michael Sintek, and Malte Kiesel. 2006. A Multilingual/Multimedia Lexicon Model for Ontologies. In *ESWC*, pages 502–513.

Philipp Cimiano, Peter Haase, Matthias Herold, Matthias Mantel, and Paul Buitelaar. 2007. LexOnto: A Model for Ontology Lexicons for Ontology-based NLP. In *Proceedings of OntoLex - From Text to Knowledge: The Lexicon/Ontology Interface (workshop at the International Semantic Web Conference)*.

Gil Francopoulo, Nuria Be, Monte George, Nicoletta Calzolari, Monica Monachini, Mandy Pet, and Claudia Soria. 2007. Lexical markup framework: ISO standard for semantic information in NLP lexicons. In *Workshop of the GLDV Working Group on Lexicography at the Biennial Spring Conference of the GLDV*.

Yue Ma, Laurent Audibert, and Adeline Nazarenko. 2009. Ontologies étendues pour l'annotation sémantique. In *20mes Journées Francophones d'Ingénierie des Connaissances (IC09)*.

Nouha Omrane, Adeline Nazarenko, Peter Rosina, Sylvie Szulman, and Christoph Westphal. 2011a. Lexicalized ontology for a business rules management platform: An automotive use case. In *RuleML@BRF*, Florida, États-Unis, November.

Nouha Omrane, Adeline Nazarenko, and Sylvie Szulman. 2011b. Les entités nommées : éléments pour la conceptualisation. In *22mes Journées Francophones d'Ingénierie des Connaissances (IC11)*.

W. Peters, M. Espinoza, E. Montiel-Ponsoda, and M. Sini. 2009. Multilingual and localization support for ontologies. Technical report, D2.4.3 Neon Project Deliverable. Technical report.

Axel Reymonet, Jérme Thomas, and Nathalie Aussenac-Gilles. 2007. Modélisation de ressources termino-ontologiques en OWL. In Francky Trichet, editor, *Journées Francophones d'Ingénierie des Connaissances (IC), Grenoble*, pages 169–180, July.